DOCUMENT RESUME

ED 388 287 IR 017 436

AUTHOR Rauterberg, M.; And Others
TITLE Automatic Sound Generation for Spherical Objects Hitting Straight Beams Based on Physical Models.
PUB DATE 94
NOTE 7p.; In: Educational Multimedia and Hypermedia, 1994. Proceedings of ED-MEDIA 94--World Conference on Educational Multimedia and Hypermedia (Vancouver, British Columbia, Canada, June 25-30, 1994); see IR 017 359.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Acoustical Environment; *Acoustics; Computer Simulation; Computer System Design; Educational Environment; *Feedback; Foreign Countries; *Human Factors Engineering; Interaction; Learning Activities; Use Studies
IDENTIFIERS *Audio Feedback

ABSTRACT

Sounds are the result of one or several interactions between one or several objects at a certain place and in a certain environment; the attributes of every interaction influence the generated sound. The following factors influence users in human/computer interaction: the organization of the learning environment, the content of the learning tasks, the temporal share of the activity on the screen in proportion to the whole activity, the content of computer support, and the user friendliness of the interactive system. An experiment that investigated the effects of audio feedback showed that the results of a database query at the user interface with individually selective acoustic feedback were as good as with a previously adjusted standard interface without any acoustic feedback. The results of a second experiment showed that sound feedback significantly improves the performance of operating and controlling the simulation system of an assembly line with 38 difference sounds. A framework concept for the description of sounds is presented in which sounds can be represented as auditory signal patterns along several descriptive dimensions of objects interacting in an environment. The methodology can be demonstrated by the falling of a spherical elastic object onto a linear elastic beam. Rather than assigning an unchangeable sound to an object or operation, which results in a synthetic tone, a sound should be calculated in real time and be context sensitive. (Contains 19 references.) (AEF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 388 287

Automatic sound generation for spherical objects hitting straight beams based on physical models

M. RAUTERBERG
Usability Laboratory, Work and Organisational Psychology Unit
Swiss Federal Institute of Technology (ETH), Nelkenstrasse 11, CH-8092 Zurich, Switzerland

M. MOTAVALLI
Swiss Federal Laboratories for Material Testing and Research (EMPA)
Uberlandstrasse 129, CH-8600 Dubendorf, Switzerland

A. DARVISHI and H. SCHAUER
Department of Computer Science (Ifi)
University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Abstract: The objective of this paper is the development of concepts, methods and a prototype for an audio frame work. This audio frame work shall describe sounds on a highly abstract semantic level. We describe every sound as the result of one or several interactions between one or several objects at a certain place and in a certain environment. The attributes of every interaction influence the generated sound. Simultaneously, the participating objects, which take part in the sound generation process, can consist of different physical conditions (states of aggregation), materials as well as their configurations. All relevant attributes have an influence on the generated sound. The hearing of sounds in everyday life is based on the perception of events and not on the perception of sounds as such. For this reason, everyday sounds are often described by the events they are based on. In this paper, a framework concept for the description of sounds is presented, in which sounds can be represented as auditory signal patterns along several descriptive dimensions of various objects interacting together in a certain environment. On the basis of the differentiation of purely physical and purely semantic descriptive dimensions, the automatic sound generation is discussed on the physical and semantic levels. Within the scope of this research project, we shall especially look for possibilities to describe the sound class 'solid objects', in particular the class of the primitive sounds 'knock' ('strike', 'hit'), because this class of sounds occurs very frequently in everyday life, the interacting objects can be easily and well described by their material characteristics and the knowledge of solid state physics can be used. As an example the falling of a spherical elastic object onto a linear elastic beam is physically and mathematically modelled, and implemented on a SGI workstation. The main parameters which influence the impact behaviour of such objects will be discussed. On the theoretical level, first a better overview and a better understanding of the capabilities, restrictions and problems of the existing instruments (tools) for the automatic generation of audio data can be anticipated.

State of the art

Many computers have a sound generator that produces a simple beep sound to indicate the errors. For a long time the only audio information for user interfaces contemporary modern workstations have signal processors and analogue digital converters and therefore sounds can be used in software. The terms that are used later in this chapter are defined as follows. *Audio signal pattern*: description of all perceptible audio signals. *Speech*: The description of all audio signals that have describable grammar structures. Music: Complex audio signal pattern that has rhythmic describable structures. *Tone*: Simple audio signal pattern with rhythmic describable structure. *Everyday sounds*: Audio signal patterns that have not been sufficiently researched to give a description and creation structure. We call the combined group of everyday sound and tone 'sound'. When we mean the assignment of a sound to an event we put the event identifier in inverted commas (for example: event = impact, sound = 'impact').

IR017436

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Gary H. Marks
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC).

Comparison of auditory and visual signal pattern

The textual representation of information is of most use when the user is familiar with the domain area and can demonstrate much experience and knowledge in that domain area. In comparison, more concrete (visual and auditory) representations of information that the user can query are of most use when the domain area is new and unknown. By comparing audio signal patterns with visual signal patterns the different advantages of each can be shown. Sounds and music can be used to improve the user's understanding of visual predecessors or can stand alone as independent sources of information. (For example: sounds as diagnostic support applied with the direction of a process simulation (Gaver, 1991).)

The parallel use of different media and the resulting parallel distribution of information, for example by simultaneously showing a predecessor through a concrete representation and its explanation through audio distribution, leads to a denser sharing of information. In this case, the user can dedicate his attention solely to the visual information, which has parallel audio support. This reduces the need to change the textual or other visual delivery and prevents the overflow of visual information (Gaver, 1991). The redundancy of information represented visually and auditory, as long as the representation of the information is realistically formed, is sensed not as disturbing, but instead it demands and increases information reception. It is important that with simultaneous information representation, that the information is harmonised together and that the different media are well synchronised.

Everyday sound perception

The perception of auditory signal patterns in everyday life can come in very different forms: a car driving by on the street, a dripping faucet, the confusion of voices from a crowd of people, opera music, a plane flying by, the buzz of a travel alarm clock, the beeping of a wristwatch, etc. All of these auditory signal patterns are divided into four categories: speech, music, sound and noise; sometimes noises and sounds are heard and grouped together. All of these categories are described sufficiently in the physical world through the mixing and superposition of different pitches, frequencies, volumes and sound duration. One of the essential differences between these categories, however, lies in their semantics: Speech serves primarily to convey information, while music and noise can have a pleasant or an unpleasant influence on the emotions. For musicians and other people who are intimate with this area, music and noise have a comparable semantic and informative character as speech does for the normal citizen. Besides from music and noise, the listener is interested next in the possibility of undisturbed, context free perception.

In contrast to music and noise, everyday sounds have a self standing characteristic; they are extremely context sensitive and event related (Gaver, 1986). Through the physical interaction of different everyday objects in 3-D space the sounds of everyday life are created and through propagation they become audible through the air. In comparison with music and noises, the semantically relevant dimension of sound lies not with the characteristic quality of the auditory signal pattern itself, but rather with the quality of the sound producing event as it respects the concerned object (Mountford, 1990).

This difference leads us to the conclusion that sounds are interpreted differently than music based upon their quality. When listening to music we are primarily interested in the effect of music on us; while when hearing everyday sounds we are interested in the quality of the sound producing object and the accompanying circumstances (e.g., surrounding conditions, events, etc.). Of course, music can be heard from the perspective of every day use; in this case the listener pays attention to the nature and tune of the instrument in use, to the tempo, to the acoustic, to the place of performance, etc. This method of listening to music is dependent upon the listener's knowledge of this domain field; only someone who is experienced with music will be able to extract all the various aspects from a piece of music. The average adult is, for the interpretation of everyday sounds, an expert with a large degree of knowledge from experience. This knowledge allows one to evaluate everyday sounds according to the following criteria for relevant information:

1) Information about the physical occurrence: we hear, if the fallen glass clinks or breaks.
2) Information about unseen structures: when knocking on the wall, we can hear if it is hollow.
3) Information about dynamic changes: when filling a wine glass, we can hear when it is full and runs over.
4) Information about abnormal conditions: we hear, when the car engine ceases to function properly and runs irregularly.
5) Information about occurrences outside of the visual field: the sound of footsteps behind us 'tells' us if someone is approaching.

Listening to everyday sounds is based upon the perception of events and not upon the perception of sounds in and of themselves. This fact becomes clear in the following example:

Illustrative example:
A pen dropped upon a piece of paper from a height of about 15 cm created a different sound than when it is dropped upon the hard surface of a desk. An altogether different sound is created when a rubber eraser is dropped upon the paper or, respectively, on the desk.

The sound created in each case of the previous example is neither a characteristic of any of the participating objects (pen, rubber eraser, sheet of paper, desk surface) nor a characteristic of the occurrence 'dropped' itself. The four different sounds in the examples are, with an observation that holds true to the reality of the situation solely determined by their respective interaction and environmental conditions. Everyday sounds are therefore due to a lack of better descriptive possibilities, often described through the underlying occurrence.

Every sound is also a result of one or more interactions between two or more objects in a definite place and in definite surroundings and can be defined as the following:

$$\text{Sound} = f(\text{objects, interaction, environment})$$

Every interaction possesses attributes that have an influence on the produced sound. At the same time the shared objects can participate in the production of sound from different aggregate conditions, materials, and even their configuration. The configurations of these materials possess attributes that also can have an influence on the produced sound.

Existing research projects

With the manufacture and adaptation in the world of interactive computer systems (e.g., multimedia applications) one must decide between the following three types of substitution: 1) *Propagation*: Here the quality of the propagation channel is important (i.e., bandwidth, etc.). 2) *Intake, saving, and reproduction*: this is actually time synchronous with visual processing to add extra significance. 3) *Automatic Production*: In this area the realistic generation of context sensitive, auditory signal pattern has the most significance.

We concentrate us in our research to the third area: the automatic production of every day sound. The existing research for audio with interactive systems will be listed. The computer manufacturer, Apple, researches the substitution of audio media for improvement of graphical user interface and presents the so called 'Auditory Icons' such as 'Earcons' (Gaver, 1986; Gaver, 1989; Gaver, 1990; Blattner, 1989). The Bell Communications Research Centre in collaboration with the Massachusetts Institute of Technology developed an 'Audio Window system' as an analogue of the 'Visual Window system' that realised one, two, and three dimensional representations of auditory source (Ludwig, 1990; Wenzel, 1991). The Olivetti Research Centre was developing a so called 'Audio Server' an analogy to the 'Window Server' with a user interface based on the Client Server Model (Binding, 1990). None of the three projects mentioned above, however, allow the description of sounds on a highly abstract semantic level.

State of own research

It is mainly the following factors in their respective forms and combinations that influence the load for humans in the man computer interaction: the organisation of the learning place and the learning place environment, the content of the learning tasks on the screen, the temporal share of the activity on the screen in proportion to the whole activity, the content of the computer support and the user friendliness of the interactive system.

Our investigations have shown that the user friendliness of a data processing system crucially influences its acceptance by its users. Three different aspects should hereby be discriminated: (1) the functionality and the amount of information, (2) the availability (response times and disturbances) and (3) the operational handling mediated by the user interface (Rauterberg, 1992).

The results of an experiment, that investigated the effects of audio feedback, showed, that the results of a database query at the user interface with an individually selective acoustic feedback were as good as with a previously adjusted standard interface without any acoustic feedback. However, if the users, who considered the acoustic feedback as useful, are compared with those, who considered it not necessary, significant performance advantage results for the persons with a positive mind about the acoustic feedback (Rauterberg et al, 1991).

The results of a second experiment with or without sound feedback in a process control environment showed, that sound feedback improves significantly the performance of operating and controlling the simulation system of an assembly line with 38 different sounds (Rauterberg & Styger, 1994).

Definition of the main research goal

From the analysis of existing models and concepts as well as empirical investigations of the pre phase a definition for the requirements for a new concept and a new model should be derived. The example of an impact (e.g., falling spherical object on the surface of a beam) shows the physical process of producing sound through the interaction of the objects 'spherical object' and 'beam' (fig. 1). The produced sound 'spherical object hits the beam' can be described by the behaviour of the vibration of the beam surface, which has been produced by the deformation resulting from the fall of the spherical object. "When an object is deformed by an external force, internal restoring forces cause a build up potential energy. When the external force is removed, the object's potential energy is transformed to kinetic energy, and it swings through its original position. The object continues to vibrate until the initial input of energy is lost" (Gaver, 1993).

The following concept should help us to find a classification to describe the sounds on a very high abstract and semantic level. The highest level in our sound hierarchy contains the class of everyday sounds that are produced through interaction of two or more objects (interacting objects). The second highest level contains the three different material states (solid, liquid, gaseous), which produce qualitatively diverse sounds. For example, if two or more solid objects interact, then they produce sounds like 'push', 'break', etc. Sounds like 'drop', 'sprinkle' are produced through interaction of different liquid objects, and sounds like 'explode', 'stream' are produced through aerodynamic interactions in the air. All of these so called 'primitive sounds' are characteristically for their classes.

Audio framework

The term 'framework' comes from the object oriented world and is understood as follows: Already programmed frames that allow the programmers to supplement the application of specific parts in a program. An audio framework should in the term of object-oriented terminology provide the basic sound classes and the possibilities to combine and manipulate them. At the moment there is no such audio framework. As an example we are going to approach the sound class 'impacts'.

Our methodology can be demonstrated by falling a linear elastic spherical object with the elasticity modules E_1, Poisson ratio ν_1, mass m and radius r_S onto a straight beam (fig. 1). The material behaviour of the beam is assumed to be isotropic and linear elastic with the elastic modules E, Poisson ratio ν, cross sectional area A = (h * b), length l and the material density ρ. First, we will ignore the damping of the system which is mainly affected by the viscoelasticity effects of the material. The behaviour of anisotropic materials such as multi-layered composite structures has been previously investigated (see for example Motavalli, 1991). The results of this investigation can be used for the further developments.

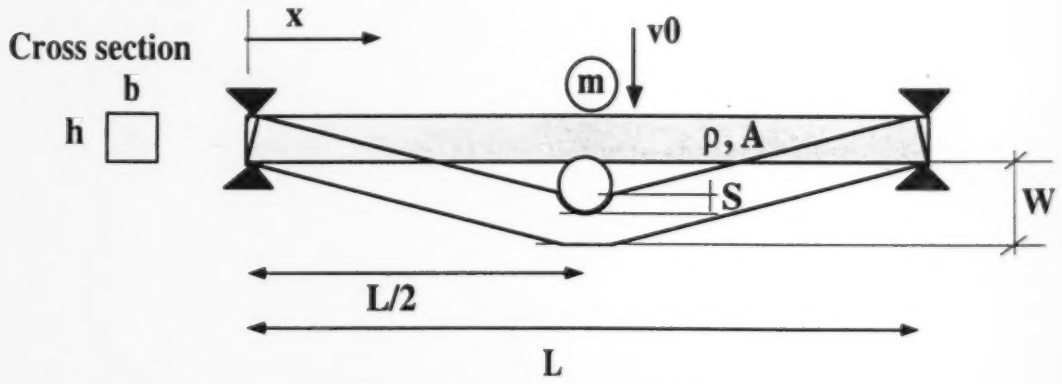

Fig. 1: Falling a mass on a straight beam

The displacement Z(t) of the centre of mass of the spherical object hitting the structure satisfies the equation of motion:

$$Z_{,tt} = -\frac{F}{m} \qquad (1)$$

where $(.)_{,tt} = d^2(.) / dt^2$ is the second time derivative, $F \geq 0$ is the resultant force acting at the contact region between the beam and the mass.

The resultant contact Force F can be connected with the maximum relative displacement S of the upper surface of the structure with respect to its middle surface according to (fig. 1) by solving the corresponding classical contact problem of linear elasticity (Hertz, 1881):

$$F = kS^{\frac{3}{2}} \tag{2}$$

where $k = \frac{4}{3}\sqrt{r_s}\left(\frac{E' E_1'}{E' + E_1'}\right)$, $E' = \frac{E}{1-\nu^2}$ and $E_1' = \frac{E_1}{1-\nu_1^2}$.

One of the most essential assumption in deriving equations (1) and (2) is that the elastodynamic effects in the region of contact are neglected. Furthermore the accuracy of the following equations is quite satisfactory as long as the radius r_S of the sphere is sufficiently large with respect to the beam thickness h (for example, $r_S/h \geq 2$).

The differential equation of the motion of the beam under distributed load q(t,x) in terms of lateral displacement W is:

$$EIW_{,xxxx} + \rho A W_{,tt} = q(t,x) \tag{3}$$

where $(.)_{,xxxx} = d^4(.)/dx^4$ and $I = (bh^3)/12$.

Taking the boundary conditions into account, this equation can be solved analytically. After some mathematical operations, the following solution can be derived:

$$W(t,x) = \sum_n \frac{2}{\rho AL}(-1)^{\frac{n-1}{2}} \sin\left(n\frac{\pi x}{L}\right)\left[\frac{1}{\omega_n}\int_0^t F(\tau)\sin(\omega_n(t-\tau))d\tau\right] \tag{4}$$

The displacement Z(t) of the centre of the sphere will be given by the sum of the two displacements S and W (fig. 1):

$$Z = S + W \tag{5}$$

The displacement W at the mid span of the beam is:

$$W(t, L/2) = \sum_{n=1,3,5,\ldots} \frac{2}{\rho AL\omega_n}\int_0^t F(\tau)\sin(\omega_n(t-\tau))d\tau] \tag{6}$$

where ω_n are the eigen frequencies of the beam.

Provided that one considers the material damping as an additional effect to the above mentioned model, the following equation can be derived instead of (4):

$$W(t, L/2) = \sum_{n=1,3,5,\ldots} \exp(-\delta_n t)\frac{2}{\rho AL\omega_n}\int_0^t F(\tau)\sin(\omega_n(t-\tau))d\tau] \tag{7}$$

where δ_n is the damping constant of the beam material.

Equations (2), (5) and (7) illustrate the dependency of the beam-surface vibration (S + W) on the following important parameter:

- material properties and geometry of the beam and
- material properties, geometry and initial velocity of the falling object.

The main advantage of the presented physical model to the existing model in (Gaver,1993) is that the material and geometrical properties of both objects are considered rather than the properties of the impacted beam only. Further investigations are needed to develop solutions for objects with other geometry.

The sound 'ball hits the straight beam' can be automatically generated using the appropriate parameters derived from the physical model of the interacting objects. Using a flat plate (Koller, 1983) and/or a non flat plate (Sayir, 1992) can cause complex mechanical equations.

The importance of an audio framework

Modern workstations (e.g., NEXT, SGI) provides sound and music kits. They also provide so called basic operations for manipulation of sound patterns in the frequency level. In addition they offer the opportunity to assign a digitised sound to a certain dialogue element, for example a button. All of these approaches assign an unchangeable sound to an object or operation. These sounds are very synthetic. It means context free and does not fulfil our expectation of the real world. Instead of that we understand a sound as a result of one or more interactions between two or more objects in a certain place and a certain environment. It means that a sound can convey more information about the sound source, it's place and environment. Therefore it should be calculated in real time and is context sensitive.

On the theoretical level a better overview and understanding of the skills and restrictions of existing tools for automatic generation of audio data are expected. The new concepts that are planned for development will give a

basic impulse to the following fields. *Multimedia*: New possibilities for interactive creation of sound producing events and actions. *Simulation , Computer supported learning*. A pedagogic target of computer support learning will clarify the processes. These processes could be physical or chemical processes. *User interfaces* (e.g., for the visually impaired): Simulations also find their applications in the future for handicapped and especially visually impaired computer users. For example, blind computer users could recognise the dialogue objects and movable pictures in a 3-D space by sounds. *CAD, Architecture*: New possibility of examining the sound isolation of rooms, walls, materials, etc. Generally, the acoustics of buildings could be changed and checked interactively.

References

Binding, C., Schmandt, C., Lantz, K. & Arons, B. (1990) Workstation audio and window-based graphics: similarities and differences. in G. Cockton (ed.) Engineering for Human-Computer Interaction. (pp. 233-247), Amsterdam: North-Holland.

Blattner, M. et al. (1989) Earcons and Icons: their structure and common design principles. *Human Computer Interaction*, 4:11-44.

Gaver, W. (1986) Auditory Icons: using sound in computer interfaces. *Human Computer Interaction*, 2, 167-177.

Gaver, W. (1989) The SonicFinder: An Interface That Uses Auditory Icons. *Human Computer Interaction*, 4, 67-94.

Gaver, W. & Smith, R. (1990) Auditory icons in large-scale collaborative environments. in D. Diaper et al. (eds.) Human Computer Interaction INTERACT'90 (pp. 735-740). Amsterdam: Elsevier.

Gaver, W. (1991) The SonicFinder: An Interface That Uses Auditory Icons. *Human Computer Interaction*, 4, 67-94.

Gaver, W., Smith, R. & O'Shea, T. (1991) Effective sounds in complex systems: the ARKola simulation. in S. Robertson, G. Olson & J. Olson (eds.), Reaching through technology CHI'91. (pp. 85-90), Reading MA: Addison-Wesley.

Gaver, W. (1993) What in the World do We Hear? An Ecological Approach to Auditory Event Perception. *Ecological Psychology*, 5(1).

Hertz, H. (1881) Über die Berührung fester elastischer Körper, Journal für die reine und angewandte Mathematik, vol. 92, pp 156-171.

Koller, M. G. (1983) Elastischer Stoss von Kugeln auf dicke Platten. Dissertation Nr. 7299, Zuerich: ETH.

Ludwig L., Pincever, N. & Cohen, M. (1990) Extending the notion of a window system to audio. *Computer*, 23(8), 66-72.

Motavalli M., (1991) Das Verhalten mehrschichtiger Verbundrohre unter Innendruck. Dissertation Nr. 9597, Zuerich: ETH.

Mountford, S. & Gaver, W. (1990) Talking and listening to computers. in B. Laurel & S. Mountford (eds.) The Art of Human-Computer Interface Design. (pp. 319-334), Reading MA: Addison-Wesley.

Rauterberg, M., Baumgartner, A., Jenny, A. & de Lange, M. (1991) Erhoehung der Transparenz durch individualisierte Soundunterstuetzung bei Datenbankrecherchen. Technical Report IS-1-91. Institut fuer Arbeitspsychologie. Zuerich: Eidgenoessische Technische Hochschule.

Rauterberg, M. (1992) An empirical comparison of menu-selection (CUI) and desktop (GUI) computer programs carried out by beginners and experts. *Behaviour and Information Technology*, 11(4), 227-236.

Rauterberg, M. & Styger, E. (1994) Positive effects of sound feedback during the operation of a plant simulator. submitted to "East-West International Conference on Human-Computer Interaction EW-HCI'94" in St. Petersburg (Russia), August 2-6, 1994.

Sayir, M. B., Haeusler, K. & Partel, M. (1992) Impact Behavior of corrugated sheets of fibre-reinforced cement. *Composites*, 23(6).

Smith, S., Bergeron R. & Grinstein, G. (1990) Stereophonic and surface sound for exploratory data analysis. in Proceedings of Computer-Human-Interaction CHI'90. New York: ACM.

Wenzel E., Wightman F. & Kistler D. (1991) Localization with non-individualized virtual acoustic display cues. in S. Robertson, G. Olson & J. Olson (eds.), Reaching through technology CHI'91. (pp. 351-359), Reading MA: Addison-Wesley.

BEST COPY AVAILABLE